DOCUMENT RESUME

ED 420 522 SE 061 561

AUTHOR Manhart, Jim J.
TITLE Gender Differences in Scientific Literacy.
PUB DATE 1998-00-00
NOTE 31p.: Paper presented at the Annual Meeting of the National Council on Measurement in Education (1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Cognitive Processes; *Concept Formation; Evaluation Methods; Grade 10; Grade 9; High Schools; Inquiry; *Knowledge Representation; Learning Strategies; *Science Education; *Scientific Literacy; Scientific Methodology; *Sex Differences
IDENTIFIERS National Science Education Standards; Nature of Science

ABSTRACT

This study investigates gender differences with regard to three factors of scientific literacy. The study involved students (N=772) in Grades 9 and 10. A 100-item multiple choice test based on National Science Education Standards was used to assess scientific literacy while gender differences were explored using analysis of variance procedures. Males tended to perform better than females on the constructs of science factor. Females tended to do better than males on the abilities necessary to do scientific inquiry factor and the social aspects of science factor. The Manel-Haenszel procedure flagged 12 items functioning differently for females and males of the same ability and all of the flagged items exhibited only small to moderate differential item functioning. (Author/DDR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 420 522

J. Manhart

2A

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Gender Differences in Scientific Literacy

Jim J. Manhart

The University of Iowa

Paper presented at the 1998 annual meeting of the

National Council on Measurement in Education

SE061501

Abstract

Gender differences with respect to three factors of scientific literacy were investigated. Participants included 772 students (387 females, 385 males) in grades 9 and 10. Based on National Science Education Standards, a 100 item multiple-choice test was constructed to assess scientific literacy. Gender differences were examined using analysis of variance procedures. Males tended to perform better than females on the constructs of science factor. Females tended to perform better than males on the abilities necessary to do scientific inquiry factor and the social aspects of science factor. The Mantel-Haenszel procedure flagged 12 items functioning differently for females and males of the same ability (8 in favor of females, 4 in favor of males). All of the flagged items exhibited only small to moderate differential item functioning.

The National Science Education Standards (National Research Council, 1996) defines scientific literacy through eight content standards. Three of the standards deal with the traditional content areas of physical science, life science, and earth/space science. The other five standards are entitled unifying concepts and processes, science as inquiry, science and technology, science in personal and social perspectives, and history and nature of science. Since the post-Sputnik reforms in science education, knowledge of science constructs and scientific methods have generally been accepted as aspects of scientific literacy. However, Standards indicates that one must also understand the social aspects of science to be considered scientifically literate today. The social aspects include not only relationships between science, technology, and society, but also the social interactions that occur among scientists as new knowledge is constructed.

Over the past few decades, a number of investigations have reported that high schools males tend to score slightly higher on science tests than females (DeBaz, 1994; O'Sullivan, Reese, & Mazzeo, 1997; Cole, 1997). However, science tests in the past have usually focused on science constructs. Although these tests typically included a few items relating to scientific methods, most did not contain items dealing with the social aspects of science. Thus, these tests were not comprehensive assessments of scientific literacy as defined in Standards. Moreover, on items dealing with scientific methods, females perform

as well or better than males (Jones, Mullis, Raizen, Weiss, & Weston, 1992; Lock, 1992; Robertson, 1987).

The purpose of the present study was to investigate gender differences with respect to the various aspects of scientific literacy. Gender differences with respect to knowledge about the social aspects of science were of particular interest, given that these aspects were not investigated in the past. Although current reforms in science education call for a reduction in the number of science constructs students need to know, it was hypothesized that males would still perform better than females on items relating to science constructs. On the other hand, in light of previous research findings, it was anticipated that females would perform as well or better than males on items dealing with scientific methods.

Methods

Participants

The sample consisted of 772 students in grades 9 and 10 who were taking a science class at four Midwestern high schools. There were 387 females and 385 males. The subjects were predominately white.

Two years of science were required at each school (physical science in grade 9, biology in grade 10). Therefore, the females and males in each grade had taken the same science courses. Three of the schools (Schools A, B, and D) were actively involved in science education reform through the Scope, Sequence, and Coordination [SS&C] site located at the University of Iowa.

SS&C is a reform project of the National Science Teachers Association. The Iowa site focuses on relationships between science, technology, and society.

Instrument

The instrument employed to assess scientific literacy consisted of 100 multiple-choice items. The multiple-choice format was selected so that scientific literacy of individual students could be comprehensively assessed in a reasonable amount of time. Since no multiple-choice test which comprehensively assesses scientific literacy was located, such a test was constructed specifically for this investigation.

The definition of scientific literacy given by National Science Education Standards (National Research Council, 1996) was the main source in creating the test specifications. The test specifications include the following eight subtests: physical science, life science, earth/space science, abilities necessary to do scientific inquiry, nature of scientific inquiry/knowledge, science as a human endeavor, science and technology, and societal perspectives. Each test item was tied to some aspect of one of the content standards. The test specifications and the relationships between the subtests and the content standards are given in the appendix. There is not a one-to-one correspondence between the subtests and the standards. The standards were purposely written to overlap in content, but mutually exclusive subtests are usually desired. Therefore, a given subtest may address more than one standard, and the aspects of a specific standard may be found on more than one subtest.

Once the test specifications were set, a number of currently available tests were examined to see if they contained any items or ideas for items that might be used on the new test. Two tests were found to be particularly useful in this regard. On the Test on Understanding Science [TOUS] (Cooley & Klopfer, 1961), about 15 items were judged to be aligned with the test specifications relating to the social aspects of science. These TOUS items served as starting points, but were rewritten to meet current standards for test items with respect to nonsexist language and readability. Items on the science test of the Iowa Tests of Educational Development [ITED], Form M (1996) were found to be an acceptable way to assess the abilities necessary to do scientific inquiry, given that the new test was to be a paper and pencil test and not include a performance assessment. Many of the items on the ITED relate to reading passages which describe a particular scientific study. These items ask students to identify the question guiding the scientific investigation, to consider issues dealing with the design of the study, and to determine what conclusions are consistent with the given data. Since Form M is a recent edition of the ITED, permission was obtained to use ITED items rather than write new items for the abilities necessary to do scientific inquiry subtest.

For the other seven subtests, new items were written and pilot tested in 16 Iowa schools during the fall of 1996. The item statistics from the pilot testing were considered when selecting items for the final version of the test. Although more items assessing the social aspects of science were desired (particularly on

the science and technology subtest and the societal perspectives subtest), several items were unfortunately not able to be used in light of poor item statistics from the pilot testing.

The test was divided into two parts of 50 questions each. This was done to allow schools the option of administering it over two days rather than on a single day. Each part was stratified by the eight subtests so the two parts would be equivalent in content. The time limit for each part was set at 40 minutes.

Confirmatory factor analysis of item parcels suggested that a three factor model was the best way to explain the covariance structure obtained from this scientific literacy test (Manhart, 1997). The model appeared to apply equally well to females and males. The three factors were labeled constructs of science, abilities necessary to do scientific inquiry, and social aspects of science. The constructs of science included the 36 items making up the physical science, life science, and earth science subtests. The abilities necessary to do scientific inquiry were assessed by the 30 items from the subtest with the same name. The social aspects of science consisted of the 34 items in the remaining four subtests (nature of scientific inquiry/knowledge, science as a human endeavor, science and technology, societal perspectives). Coefficient alpha was .80 for constructs of science, .86 for abilities necessary to do scientific inquiry, and .84 for social aspects of science.

Procedures

Since during the fall of 1996 each school gave the science test of the ITED (Form L, 1993), the data from that test was examined to see how the students who participated in this study compare with other students in the nation. The percentages of students in this study with scores on the ITED science test between selected percentile intervals were calculated and compared with the national percentages.

The scientific literacy test was administered at each school during February or March of 1997. Schools B, C, and D administered the test in a single session; School A gave the test over a two day period. Items were scored dichotomously; a response matching the key received a score of one and all other responses or no response received a score of zero. Raw scores for each of the three factors of scientific literacy were computed by summing the scores for the individual items. The mean and standard deviation were calculated for each factor. Correlations between the factors were also computed.

A separate analysis of variance [ANOVA] was performed for each factor of scientific literacy. Since gender differences could possibly vary as a function of school and grade, three-way ANOVA (School X Grade X Gender) were employed. All three effects were considered fixed. The analysis was performed using the General Linear Model [GLM]. In GLM, the sums of squares depend on the order in which they are calculated. Type I sum of squares analysis was utilized. The sums of squares for school, grade, and school by grade interaction

were calculated first, and then the sums of squares involving gender were calculated. All statistical tests were conducted at the .05 level.

Gender differences at the item level were also explored. Differential Item Functioning [DIF] for females and males of the same ability was examined using the Mantel-Haenszel procedure. Equal ability females and males were identified based on the factor raw score of the item being tested for DIF.

Conventions established by Educational Testing Services [ETS] were followed. These conventions include the use of the Mantel-Haenszel delta difference [MH D-DIF] statistic (Dorans & Holland, 1993). Positive MH D-DIF values indicate items differentially easier for females and negative MH D-DIF values indicate items differentially easier for males.

ETS also classifies items into three categories (A, B, C) to indicate the amount of DIF. These DIF categories are based on both the statistical significance and the absolute value of MH D-DIF (Zieky, 1993), as documented in Table 1. The DIF of items in category A is not considered to be of practical significance. Items in category B are viewed as having small to moderate DIF, while those in category C are viewed as having moderate to large DIF. In the present investigation, items falling into categories B and C were examined for possible gender bias and for any content commonalities.

Results and Discussion

The number of participants by school, grade, and gender is documented in Table 2. For each grade, the percentages of students scoring within selected

percentile intervals on the ITED science test were calculated and compared with the national percentages. As shown in Table 3, the vast majority of students participating in this study scored above the national median. Thus, the results reported for this investigation may not generalize to students at a lower achievement level.

The means and standard deviations for the raw scores of the three factors of scientific literacy are reported in Table 4. The constructs of science factor was somewhat more difficult than the abilities necessary to do scientific inquiry factor and the social aspects of science factor. The correlations between the raw scores for the three factors are shown in Table 5. The correlations were .63, .71, and .76.

ANOVA results for the three factors are shown in Table 6. Although only the effects involving gender are of concern here, for completeness all effects are included in the table. The effects are listed in the order in which the sums of squares were calculated. The triple interaction (School X Grade X Gender) was not significant for any of the three factors. Of the double interactions involving gender, the school by gender interaction was significant at the .05 level for all three factors. This suggests that gender differences should be inspected at the school level, rather than looking at the gender main effect where all schools are combined. Since the grade by gender interaction was not significant for any of the factors, there is no reason to consider the grades separately.

Table 7 contains the means for females and males by school for each of the three factors. These means are graphed in Figure 1. For the constructs of science, males scored significantly higher than females at two schools. Females scored higher than males at the other two schools, but these gender differences were not significant. Thus, there was a tendency for males to score higher than females on the constructs of science.

The constructs of science factor was assessed by three subtests. The subtest means for females and males were examined by school to note any trends. For the physical science and earth/space science subtests, the mean for males was usually higher than the mean for females. However, for the life science test, no consistent gender difference was found across the four schools. These observations are consistent with previous research reporting that higher achievement in science by males is found more in the physical sciences than in the life sciences (Johnson, 1987; Jones et al., 1992; Keeves & Kotte, 1992).

For the abilities necessary to do scientific inquiry, females scored significantly higher than males at two of the schools. Males scored higher than females at the other two schools, but these gender differences were not significant. Thus, there was a tendency for females to score higher than males on the abilities necessary to do scientific inquiry. These findings are in line with previous research reporting females do as well or better than males on scientific methods (Jones et al., 1992; Lock, 1992; Robertson, 1987).

The social aspects of science exhibited the most consistent gender differences of the three factors. At all four schools, females scored higher than males on the social aspects of science. The gender difference at two of the schools reached statistical significance. Thus, there was a tendency for females to score higher than males on the social aspects of science.

The social aspects of science factor was assessed using four subtests. The subtest means for females and males were examined by school to note any trends. At all four schools, the mean for females was higher than the mean for males on the nature of scientific inquiry/knowledge subtest, the science as a human endeavor subtest, and the societal perspectives subtest. On the science and technology subtest no consistent gender difference was found.

At the item level, the Mantel-Haenszel procedure flagged 12 of the 100 items as functioning differently for females and males of the same ability (8 in favor of females, 4 in favor of males). These items are documented in Table 8. All of the flagged items fell into category B (small to moderate DIF) of the ETS classification scheme.

Two constructs of science items were flagged for DIF. An item dealing with the relationship between electricity and magnetism was flagged in favor of males. An item dealing with animal behavior being dependent on both heredity and environment was flagged in favor of females.

The flagged items for the abilities necessary to do scientific inquiry tended to have larger MH D-DIF values than the flagged items for the other two factors.

Three of the items in favor of females related to a reading passage dealing with variables important in the bonding of infant monkeys with their mothers. The remaining item in favor of females dealt with the undesirability of stirring a water and an air sample in a particular experiment. The two items functioning in favor of males dealt with the effect of surface area on evaporation and reading a graph to determine when two materials were at the same temperature.

With respect to the social aspects of science, four items were flagged for DIF. Males performed better than females on an item asking what would not be considered part of a scientific investigation. Three items in which females performed better than males dealt with the evolution of a scientific theory, ethical treatment of human subjects, and the role of scientists in public policy decisions.

Items that are flagged for DIF are not necessarily gender biased. Items can be judged by experts as being content appropriate and still exhibit DIF. The items used in this investigation were all tied to some aspect of Standards. Therefore, the gender differences observed in this investigation were probably not due to biased test items. Moreover, all of the flagged items only exhibited small to moderate DIF.

The flagged items do not appear to have much content in common. One possible exception is with respect to animal behavior. The item dealing with animal behavior being dependent on heredity and environment, and three items tied to the reading passage about infant monkeys bonding to their mother, were flagged in favor of females. Thus, females may have a better understanding of

animal behavior than males. Females may also have been more interested, and thus more engaged, with the reading passage than males.

Implications and Limitations

If future tests for the high school level are constructed based on the definition of scientific literacy given in Standards, gender differences on overall science achievement in favor of males may disappear. Science tests in the past have focused on assessing knowledge of science constructs. Such tests show small, but consistent gender differences in favor of males. However, the results of the present investigation suggest that females tend to do better than males on abilities necessary to do scientific inquiry and social aspects of science. Thus, if science tests in the future place more emphasis on these other two aspects of scientific literacy, then gender difference in overall science achievement may disappear or may even change in favor of females.

In the present investigation, equal number of items were included for physical science, life science, and earth science. This decision was based on the fact that each is a content standard found in Standards. However, Standards does not indicate what emphasis should be placed on each standard, but only that all of the standards must be included in science education. Based on opportunity to learn considerations, a case may be made for having just as many life science items as physical science and earth/space science items combined. Tests of scientific literacy are intended for all students, and therefore to be fair should only be based on required science courses. A typical sequence for the

last four years of required science is as follows: life science in grade 7, earth/space science in grade 8, physical science in grade 9, and biology in grade 10. In light of these required science classes, life science should be assessed with the same emphasis as physical/earth/space science. With such an increase in emphasis on life science, the tendency for males to score higher than females on constructs of science would likely lessen or even be eliminated.

The results relating to gender differences also have implications for classroom teachers attempting to implement the science content standards. Teachers may need to give special attention and encouragement to males when social aspects of science are considered, whereas females may need such special treatment when studying constructs of science in the physical and earth/space sciences. Items flagged for DIF may not be gender biased, but may have implications for classroom instruction. For instance, the DIF analysis in this investigation suggested that questions relating to animal behavior may be somewhat easier for females than males. Males may need more, and possibly different, instruction in this topic than females.

The schools that volunteered to participate in this study were not a representative sample of high schools across the United States. All four schools were located in the Midwest. Students at these schools were predominately white. The vast majority of the students scored above the national median on the ITED science test. The results may generalize to other predominately white Midwestern high schools. However, whether the results generalize to other

populations is more questionable. On the other hand, the results for the constructs of science and abilities necessary to do scientific inquiry were similar to those reported in previous investigations conducted on representative samples in the United States and in other countries (e.g., Jones et. al, 1992; Keeves & Kotte, 1992). Thus, there is reason to expect that the gender difference in favor of females on the social aspects of science may generalize to other populations.

The scientific literacy test employed in this study used the multiple-choice format exclusively. Whether the same gender differences would be found using constructed response, performance assessments, or a mixture of formats remains to be seen. Future studies might investigate whether item format has an impact in the assessment of scientific literacy.

References

Cole, N. S. (1997). The ETS gender study: How females and males perform in educational settings. Princeton, NJ: Educational Testing Service.

Cooley, W. W., & Klopfer, L. E. (1961). Test on understanding science (Form W). Princeton, NJ: Educational Testing Service.

DeBaz, T. P. (1994). A meta-analysis of the relationship between students' characteristics and achievement and attitudes toward science (Doctoral dissertation, Ohio State University, 1994). ProQuest - Dissertation Abstracts [CD-ROM]. Item: 9516981

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), Differential item functioning (pp. 35-66). Hillsdale, NJ: Lawrence Erlbaum Associates.

Iowa tests of educational development (Form L: Levels 15, 16, 17/18). (1993). Chicago: Riverside Publishing.

Iowa tests of educational development (Form M: Levels 15, 16, 17/18). (1996). Chicago: Riverside Publishing.

Johnson, S. (1987). Gender differences in science: Parallels in interest, experience and performance. International Journal of Science Education, 9, 467-481.

Jones, L. R., Mullis, I. V., Raizen, S. A., Weiss, I. R., & Weston, E. A. (1992). The 1990 science report card: NAEP's assessment of fourth, eighth and twelfth graders. Washington, D. C.: National Center for Education Statistics.

Keeves, J. P., & Kotte, D. (1992). Disparities between the sexes in science education: 1970-84. In J. P. Keeves (Ed.), The IEA study of science III: Changes in science education and achievement: 1970 to 1984 (pp. 141-164). Elmsford, New York: Pergamon Press.

Lock, R. (1992). Gender and practical skill performance in science. Journal of Research in Science Teaching, 29, 227-241.

Manhart, J. J. (1997). Scientific Literacy: Factor Structure and Gender Differences. Unpublished doctoral dissertation, The University of Iowa, Iowa City.

National Research Council. (1996) National science education standards. Washington, DC: National Academy Press.

O'Sullivan, C. Y., Reese, C. M., & Mazzeo, J. (1997). NAEP 1996 science report card for the nation and the states. Washington, D. C.: National Center for Education Statistics.

Robertson, I. J. (1987). Girls and boys and practical science. International Journal of Science Education, 9, 505-518.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), Differential item functioning (pp. 337-347). Hillsdale, NJ: Lawrence Erlbaum Associates.

Appendix:
Test Specifications

Physical Science Subtest [PS]

12 items based on the Physical Science standard

- Structure of atoms (includes nuclear changes)
- Structure, properties, and changes in properties of matter
- Chemical reactions
- Motion and forces (gravitational, electrical, nuclear)
- Transfer/conservation of energy and increase in disorder
- Interactions of energy and matter (includes waves)

Life Science Subtest [LS]

12 items based on the Life Science standard

- Structure and function in living systems (includes cells and biochemistry)
- Reproduction and heredity (including molecular basis)
- Biological evolution, diversity, and adaptations of organisms (includes classification)
- Populations, ecosystems, and interdependence of organisms
- Matter, energy, and organization in living systems
- Regulation and behavior of organisms

Earth/Space Science Subtest [ES]

12 items based on the Earth and Space Science standard

- Earth and moon in the solar system
- Structure of the earth system
- Energy in the earth system
- Geochemical cycles
- Origin and evolution of the earth system and solar system
- Origin and evolution of the universe

Appendix (cont.)

Abilities necessary to do Scientific Inquiry Subtest [ASI]

30 items based on the Science as Inquiry standard

- Identify the questions (purposes) guiding specific scientific investigations
- Design scientific investigations (includes clarifying assumptions and variables)
- Develop descriptions, explanations, models, and predictions using evidence, logic, and mathematics

Nature of Scientific Inquiry/Knowledge Subtest [NSI]

12 items based mainly on the Science as Inquiry standard and the History and Nature of Science standard

- What science can investigate; What science cannot investigate
- Criteria for acceptance as a scientific fact or explanation; How science differs from other ways of knowing
- Hypotheses and conceptual principles influence the design and interpretation of scientific investigations
- How scientific knowledge develops (includes use of math, models, and theories)

Science as a Human Endeavor Subtest [SHE]

12 items based mainly on the History and Nature of Science standard and the Science in Personal and Social Perspectives standard

- People from various cultures and both genders contribute to science today. In the past, racial minorities and women have been left out of science.
- Scientists are in communication with each other regardless of where and how they do scientific investigations.
- Scientists have some assumptions, attitudes, and values in common. Examples include the universality of the laws of nature, skepticism, and accurate reporting of data, respectively.
- Scientists follow ethical standards for scientific investigations. Peer review helps to ensure that scientists are honest and ethical.
- Scientists can be biased, for they are affected by their cultural and personal ways of viewing the world.
- Different scientific fields and different types of scientific inquiry require different abilities and skills. However, there are no sharp divisions between the various scientific fields and some scientists work on the border between scientific fields (e.g., biochemists).

Appendix (cont.)

Science and Technology Subtest [ST]

5 items based mainly on the Science and Technology standard and the Science as Inquiry standard

- Differences between science and technology
- Interactions between science and technology
- Constraints, tradeoffs, and consequences of technological designs

Societal Perspectives Subtest [SP]

5 items based on the Science in Personal and Social Perspectives standard

- Impact of science and technology on society
- Impact of society on science and technology
- Science and technology alone do not yield final answers on matters of public concern.

Table 1. DIF Categories Used at ETS

Category	MH D-DIF
A	not significantly different than 0 OR absolute value < 1
B	significantly different than 0 AND (1 ≤ absolute value < 1.5 OR absolute value ≥ 1.5 but not significantly > 1)
C	absolute value ≥ 1.5 AND absolute value significantly > 1

Note. significance is at the .05 level

Table 2. Number of Participants by School, Grade, and Gender

	School A	School B	School C	School D	All Schools
Grade 9	44	85	124	109	362
Females	27	49	48	54	178
Males	17	36	76	55	184
Grade 10	70	100	117	123	410
Females	35	55	58	61	209
Males	35	45	59	62	201
Both Grades	114	185	241	232	772
Females	62	104	106	115	387
Males	52	81	135	117	385

Table 3. Percent of Students Scoring Within Selected Percentile Intervals on ITED Science Test

Percentile Interval	% Students Nationally	% Students in this Study	
		Grade 9	Grade 10
90-99	10	36	37
75-90	15	23	23
50-75	25	27	28
25-50	25	11	10
10-25	15	2	2
1-10	10	1	0
% Above National Median		86	88
% Below National Median		14	12

Table 4. Means and Standard Deviations for Three Factors of Scientific Literacy

Factor	Number of Items	Mean	Standard Deviation
1	36	15.68	5.91
2	30	17.82	6.30
3	34	19.33	6.32

Note. Based on raw scores for each factor

Factor 1 = Constructs of Science
Factor 2 = Abilities necessary to do Scientific Inquiry
Factor 3 = Social Aspects of Science

Table 5. Correlations between Three Factors of Scientific Literacy

	Factor 1	Factor 2	Factor 3
Factor 1	--		
Factor 2	.71	--	
Factor 3	.63	.76	--

Note. Based on raw scores for each factor

Factor 1 = Constructs of Science
Factor 2 = Abilities necessary to do Scientific Inquiry
Factor 3 = Social Aspects of Science

Table 6. ANOVA of Raw Scores for Three Factors of Scientific Literacy

Factor 1 - Constructs of Science

Source	df	F
School	3	36.31*
Grade	1	6.37*
School X Grade	3	0.83
Gender	1	11.18*
School X Gender	3	4.75*
Grade X Gender	1	0.21
School X Grade X Gender	3	0.99
error	756	(29.81)

Factor 2 - Abilities necessary to do Scientific Inquiry

Source	df	F
School	3	40.03*
Grade	1	4.19*
School X Grade	3	1.73
Gender	1	2.48
School X Gender	3	4.33*
Grade X Gender	1	0.14
School X Grade X Gender	3	0.71
error	756	(33.91)

Factor 3 - Social Aspects of Science

Source	df	F
School	3	43.59*
Grade	1	15.46*
School X Grade	3	0.98
Gender	1	16.62*
School X Gender	3	3.15*
Grade X Gender	1	0.01
School X Grade X Gender	3	0.16
error	756	(33.03)

Note. Values in parentheses are mean square errors
*$p < .05$

Table 7. Raw Score Means and Standard Deviations by Gender for Three Factors of Scientific Literacy

Factor 1 - Constructs of Science (36 items)

	School A	School B*	School C*	School D
Females				
Mean	15.10	13.85	17.11	13.70
SD	4.68	5.10	5.79	4.79
Males				
Mean	14.69	16.43	19.78	13.45
SD	5.50	5.74	6.14	5.53

Factor 2 - Abilities necessary to do Scientific Inquiry (30 items)

	School A*	School B	School C	School D*
Females				
Mean	19.18	17.60	20.31	15.83
SD	5.75	5.45	5.81	5.29
Males				
Mean	16.02	18.00	21.07	13.91
SD	6.38	6.44	5.58	6.32

Factor 3 - Social Aspects of Science (34 items)

	School A*	School B	School C	School D*
Females				
Mean	21.81	19.26	22.72	17.50
SD	4.86	6.01	5.34	5.99
Males				
Mean	17.83	18.64	21.87	15.02
SD	5.62	7.05	5.53	5.61

*$p < .05$

Table 8. Differential Item Functioning

Factor 1 - Constructs of Science

Item	MH D-DIF	In favor of	DIF Category
PS-8	-1.07	males	B
LS-12	1.10	females	B

Factor 2 - Abilities necessary to do Scientific Inquiry

Item	MH D-DIF	In favor of	DIF Category
ASI-3	-1.55	males	B
ASI-9	-1.41	males	B
ASI-12	1.03	females	B
ASI-16	1.49	females	B
ASI-18	1.24	females	B
ASI-23	1.29	females	B

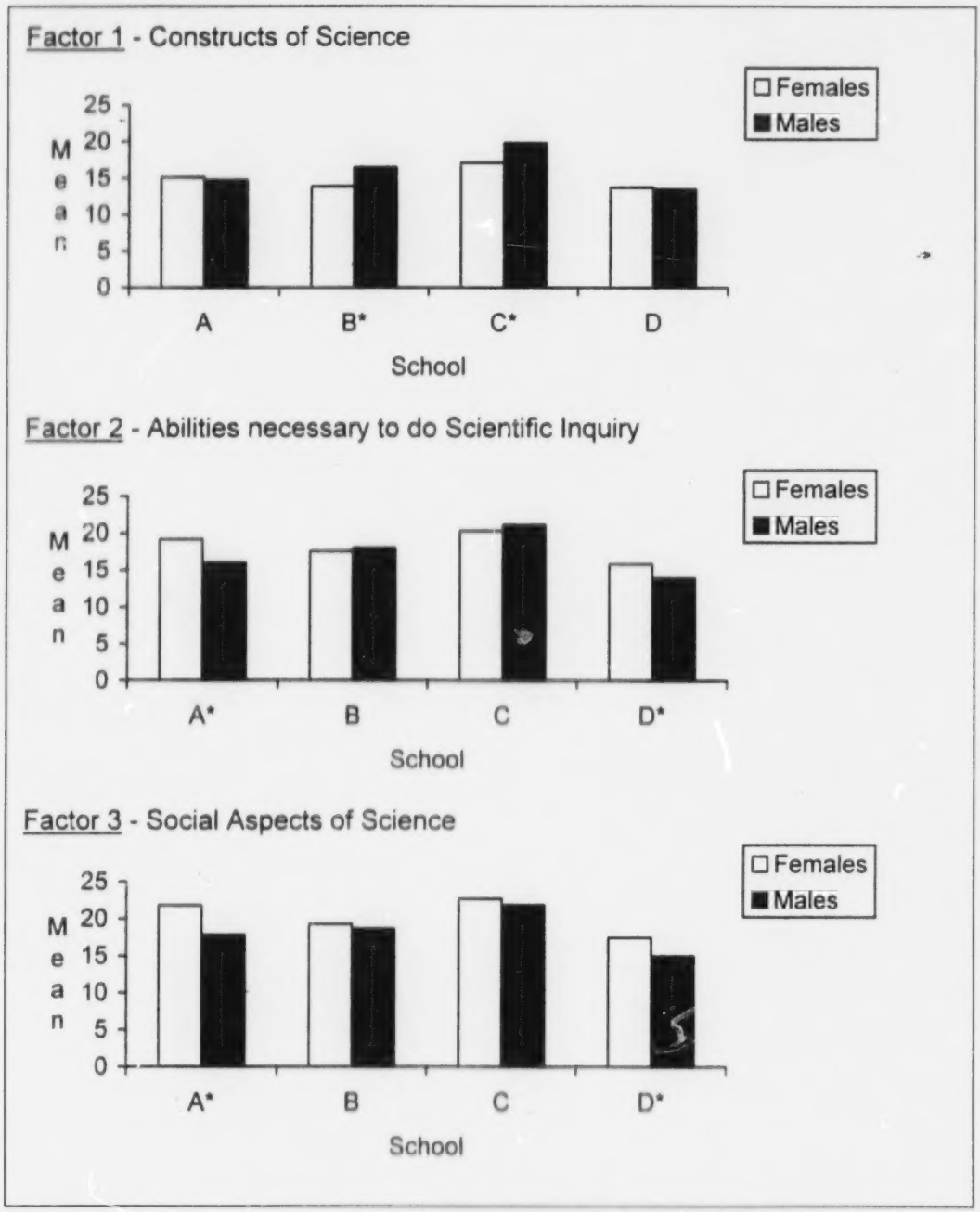
Factor 3 - Social Aspects of Science

Item	MH D-DIF	In favor of	DIF Category
NSI-2	-1.04	males	B
NSI-6	1.16	females	B
SHE-4	1.08	females	B
SP-5	1.19	females	B

Note. Subtest Abbreviations

PS = Physical Science
LS = Life Science
ES = Earth/Space Science
ASI = Abilities necessary to do Scientific Inquiry
NSI = Nature of Scientific Inquiry/Knowledge
SHE = Science as a Human Endeavor
ST = Science and Technology
SP = Societal Perspectives

Figure 1. Graphs of Raw Score Means by Gender for Three Factors of Scientific Literacy



*p < .05